

CLUSTERING VON DRUCKLUFTVOLUMENSTRÖMEN ZUR ERKENNUNG VON ANOMALIEN MIT MASCHINELLEM LERNEN

Christian DIEROLF¹, Alexander SAUER²

Inhalt

Druckluft wird in Unternehmen in verschiedenen Bereichen angewendet, dazu zählen: (i.) Antreiben von Druckluftwerkzeuge und -motoren; (ii.) Transportieren in pneumatische Förderanlagen; (iii.) Spannen, klemmen und pressen in pneumatischen Haltevorrichtungen und Pressen; (iv.) Spritzen in Spritzgussmaschinen; füllen von Nahrungsmittel; (v.) Verpacken in Verpackungsmaschinen oder Sperrluft bei Werkzeugmaschinen. [1] Für die Erzeugung von Druckluft werden in Deutschland 3 % des nationalen Gesamtstrombedarfs eingesetzt. In Industriebetrieben ist Druckluft eine teure Energieform mit jährlich verursachenden Kosten von fast 1,5 Mrd. EUR. [2] [3] [4] Unkontrolliert ausströmende Druckluft verursacht in Deutschland bis zu einem Drittel des gesamten Druckluftverbrauchs. [5]

Zur Erkennung der unkontrolliert ausströmenden Druckluft als Anomalie im Gesamtdruckluftsystem werden in der Untersuchung Druckluftvolumenströme mit Hilfe von Clustering-Algorithmen basierend auf maschinellem Lernen gruppiert. Bis auf die gewünschte Anzahl zu findender Cluster ist vorab kein zusätzliches Wissen über das Druckluftsystem erforderlich. Eine Anomalie ist etwas Besonderes. Sie ist die Abweichung von einem erwarteten Verhalten, Muster oder einer Struktur. Maschinelles Lernen generiert Wissen aus Erfahrung oder im Anwendungsfall aus historischen Messdaten. Es existieren bessere Einsatzmöglichkeiten der menschlichen Intelligenz, als die Rund-um-die-Uhr Überwachung von Datenreihen, daher wird in der Arbeit maschinelles Lernen verwendet. Beim unüberwachten Lernen, wozu Clustering zählt, werden in repräsentativen Messdaten Gruppen erkannt. Die Zuordnung der Daten zu den Gruppen nicht explizit vorgegeben.

Messung und Methode

In der Arbeit ist der Gesamtvolumenstrom einer Druckluftanlage analysiert, bestehend aus sechs Kompressoren mit insgesamt 130 kW Nennleistung und einem Druckluftspeicher. Im Zwei-Schicht-Betrieb des Unternehmens sind im Messzeitraum von drei Monaten Messdaten im Messintervall von 15 Minuten aufgezeichnet. Zunächst sind die Messdaten in 163 Tagesmessungen unterteilt, die jeweils aus einem Merkmalsatz von 72 Merkmalen der 24 Stundenwerte bestehen. Merkmale sind die Werte aus dem Stundenmittel-, Stundenminimum- und Stundenmaximum. Zur Darstellung im zwei- und dreidimensionalen Raum sowie zur Reduktion der Merkmale hinsichtlich ihrer informationsverdichteten Hauptkomponenten ist die Hauptkomponentenanalyse für zwei und drei Hauptkomponenten durchgeführt [6]. Anschließend gruppiert ein Clustering-Algorithmus zwei oder drei Hauptkomponenten der Merkmalsätze zu Clustern wie folgt: Zufällige Mittelpunkte mit der Anzahl der definierte Cluster werden gewählt. Jeder Punkt aus Abbildung 1, was dem jeweiligen Merkmalsatz entspricht, wird dem am nächsten liegenden Cluster-Mittelpunkt zugeordnet. Das geometrische Zentrum der zugeordneten Punkte wird bestimmt und der Mittelpunkt dorthin verschoben. Nach anfänglich zufälligen Wahl der Mittelpunkte werden die weiteren Schritte so lange wiederholt, bis sich die Mittelpunkte nicht mehr wesentlich verändern. [7]

Ergebnisse

Mit der verwendeten Methode angewendet auf die Messdaten ist es möglich, Anomalien durch Clustering mit unüberwachten maschinellem Lernen zu gruppieren und anschließend im Signalverlauf des Volumenstroms tageweise zu interpretieren. Das Clustering der zwei Hauptkomponenten für eine vorgegebene Anzahl ist in Abbildung 1 dargestellt.

¹ Fraunhofer Institut für Produktionstechnik und Automatisierung, Nobelstraße 12, 70569 Stuttgart, Deutschland, +49 711 9703650, christian.dierolf@ipa.fraunhofer.de, www.ipa.fraunhofer.de

² Universität Stuttgart, Institut für Energieeffizienz in der Produktion, Nobelstraße 12, 70569 Stuttgart, Deutschland, +49 711 9703600, alexander.sauer@eep.uni-stuttgart.de, www.eep.uni-stuttgart.de

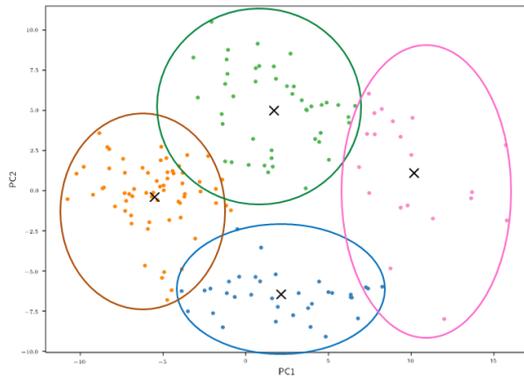


Abbildung 1: Clustering der zwei Hauptkomponenten (PC1 und PC2) der Merkmalsätze für vier Cluster

Zwei der vier Cluster sind charakteristische Volumenstromprofile für Montag und Samstag, welche in Abbildung 2 und Abbildung 3 eingetragen sind. Die beiden nicht dargestellten Cluster sind Dienstag bis Freitag und Sonntag. Weicht ein Tagesprofil einem vorgegebenen Toleranzband ab, ist diese im Signalverlauf als Anomalie erkennbar. Die Erhöhung der Anzahl von vier auf neun zu gruppierender Cluster führt bei der Untersuchung dazu, dass Anomalien durch unterschiedliche Cluster wie Samstag normal und anormal anstelle eines differenzierten Signalverlaufs durch den Algorithmus gruppiert werden.

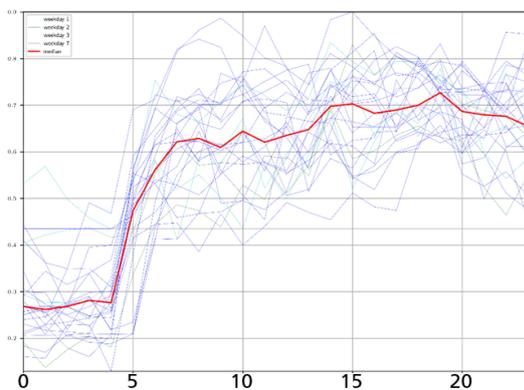


Abbildung 2: Charakteristisches Profil des Volumenstrom für Montag mit Median (rot)

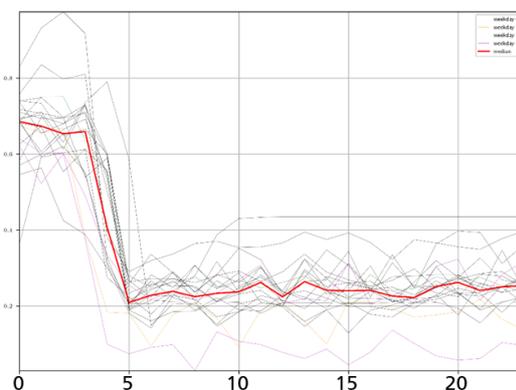


Abbildung 3: Charakteristisches Profil des Volumenstrom für Samstag mit Median (rot)

Charakteristische Tagesprofile können zur Kosteneinsparung verwendet werden indem Sie als Bedarfsprognosen für Folgetage, Monitoring von Betriebszeiten wie Effizienzmaßnahme zur Grundlastsenkung oder Identifikation wertschöpfender und nicht wertschöpfender Zeiten wie Sonntage eingesetzt werden. Eine detailliertere Unterteilung in Betriebszustände ist möglich.

Das Clustering kann durch zusätzliche Kontextinformation verbessert werden, indem Untermessungen wie Einzelmessung der Kompressoren installiert, charakteristische Stundenprofile pro Tag eingeführt oder die Messdaten mit der Produktionsplanung oder –auslastung fusioniert werden.

Referenzen

- [1] R. Gloor, „Energieeinsparungen bei Druckluftanlagen in der Schweiz - Programm Elektrizität Forschungsprojekt,“ Gloor Engineering, Sufers, 2000.
- [2] B. L. f. Umweltschutz, „Effiziente Druckluftsysteme,“ Bayerisches Landesamt für Umweltschutz, Augsburg, 2004.
- [3] Umweltbundesamt, „Entwicklung des Stromverbrauchs nach Sektoren,“ Umweltbundesamt, 2018.
- [4] Statista GmbH, „Nettostromverbrauch in Deutschland in den Jahren 1991 bis 2018 (in Terawattstunden). Statista,“ Statista GmbH, 2019.
- [5] Zentrale Koordinierungsstelle KEFF, „KEFFIZIENZ-LEITFADEN DRUCKLUFT,“ Zentrale Koordinierungsstelle KEFF, Stuttgart, 2017.
- [6] L. Smith, „A tutorial on Principal Components Analysis,“ Februar 2002. [Online]. Available: http://www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf. [Zugriff am 31 Oktober 2019].
- [7] D. Arthur und S. Vassilvitski, „k-means++: The advantages of careful seeding,“ in Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms, Society for Industrial and Applied Mathematics , 2007.